Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors & Ergonomics Society. 2001.

IMPROVING INCIDENT REPORTS USING A SCHEMATIC RECALL AID: THE CRITICAL EVENT REPORTING TOOL (CERT)

Terry L. von Thaden and Douglas A. Wiegmann
University of Illinois at Urbana-Champaign

Incident reports are intended to help identify problems in safety critical systems and aid in preventing subsequent accidents, thus improving safety. Unfortunately, incident-reporting forms are generally unstructured, providing little guidance to the reporter on how to describe the critical events of an incident. As a result, most reports contain information on *what* happened, as opposed to *why* an incident happened, and hence make the identification of intervention and prevention strategies onerous. The purpose of the present study was to help remedy this situation by developing and testing a tool, coined the Critical Event Recall Tool (CERT) for improving the quality of information contained in narrative reports of incidents. Results of this study indicate that CERT improved the analytical content of general aviation pilots' reports of a critical incident that occurred during a simulated cross-country flight and that pilots who used CERT generally felt that the tool was easy to use and understand.

INTRODUCTION

Incident reports aid in identifying problems within a system before accidents occur. As noted by Heinrich (1959), incidents are precursors to accidents, consisting in orders of magnitude significantly higher than accidents. Incidents indicate the presence of problems in systems that if left unresolved, have the potential to result in an accident. For this reason, incident reporting has long been utilized in the aviation realm as a learning tool for proactively analyzing and treating unsafe conditions and actions before they become accidents, subsequently improving safety. The most widely known incident reporting system in aviation is the National Aeronautical and Space Administration's (NASA) Aviation Safety Reporting System (ASRS). Anonymous incident reporting systems, such as the ASRS, are crucial to accident prevention, as noted by Connell (1999) because incident reports often provide information about safety hazards that is frequently lost or not obtained from post-accident investigation due to crewmember fatality.

Unfortunately, most incident reporting systems do not gather rich enough event data to understand the reasons why an unsafe event occurred. In fact, the narrative portion of most incident reporting forms are free-format essays that provide little guidance to the reporter on how to describe the critical events of an incident. As a result, most reports contain information on *what* happened, as opposed to *why* an incident happened. Consequently, identifying effective intervention strategies is often onerous.

To help remedy this situation, a tool for improving the type of information reported within the narrative portion of incident reports has recently been developed. This tool, coined the Critical Event Reporting Tool (CERT), draws upon previous work in critical incident reporting techniques (Klein et al., 1989; Militello & Hutton, 1998) and schematic organizers (Wiegmann, D, et al, 1992; Wiegmann, N, 1992) to help generate higher quality information from incident reports. Specifically, CERT functions as a pre-organizer that reporters can use for recalling event-related information prior to writing an essay. In particular, CERT was designed to encourage users to think of *why* an incident happened and the factors that affected their actions during an incident, in addition to describing the equally important *what* events that occurred during an incident.

Information Recall

As chronicled by Ericsson and Simon (1980), the process of recalling an event is limited by the capacity of Short Term Memory (STM), where only the most recently attended-to information is directly reachable. A portion of the contents of STM, however, is often instantiated in Long Term Memory (LTM) before it is lost from STM. It is this portion that can, at a later date, be retrieved from LTM. According to Tversky and Kahneman (1973) who investigated the transitional process from event to retrospective description, events that are readily recalled are judged to be representative and frequent, yet lead to large errors in estimation.

The type of information recalled may also be a function of one's expertise or experience with a given task. For example, Hall, Gott and Pokomy (1995) noted that experts are able to describe intricate aspects of their problem solving procedures, yet sometimes provide limited insight into the principles employed, failing to establish the relationship between their domain knowledge and the strategies used to solve the problems. Presumably, experts possess elaborate cognitive schemas of the domain, which improve problem solving yet make their knowledge more implicit than explicit. Novices, on the other hand, tend to report what is immediately available to their awareness. When pressed for an explanation, novices generate random reasoning. Given their lack of understanding of the domain, novices are unable to think in terms of configuring effective plans for efficient analysis.

Knowledge Elicitation

The process of how to best elicit knowledge is multifaceted and must be strategically organized to acquire the intricate knowledge structures of experts, while at the same time, eliciting thorough procedural explanations from novices. Diaper (1989) outlines a systematic elicitation process in which the elicitor devises a system that ensures all the relevant material is identified and gathered regardless of expertise. One such method is careful goal decomposition, in which the solving of a problem is broken down into subgoals or subtasks until they can be discussed in some detail that shows the interrelation of the higher and lower level corollaries of the problem.

Cognitive Task Analysis (CTA) is another method developed to elicit an operator's strategies and decisions by analyzing the various steps involved in an operator's task performance. One such CTA technique is the Critical Decision Method (CDM; Klein, 1993) in which an interviewer uses cognitive probes to understand the processes underlying decisions made by an interviewee when describing a non-routine event. Cognitive probes provide a means to focus on key decisions, cues, and options used during a critical event. This allows the interviewer to uncover a variety of aspects of the operator's decision processes, including why certain choices were made at key points in the event, as opposed to other courses of action, what aided the decision process and what might have been done had the scenario been different. CDM also delineates the differences between novice and expert performance of a task by noting the cues and inferences experts might make when performing a task and comparing these strategies to those of novices. Organizing an expert's representation of how he or she relates to a complex system allows interventions to be developed that facilitate performance and safety.

A variety of other knowledge elicitation techniques also exist, including cognitive interviewing (Loftus et al., 1989; Geiselman & Fisher, 1989; Memon & Bull, 1991), critical incident reporting techniques (Kirwan & Ainsworth, 1992; Flanagan, 1954), and even hypnosis (Hiland & Dzieszkowski, 1984). However, all of these require one-on-one interviews conducted by "knowledge extraction" experts and often require a considerable amount of time to employ. Consequently, none are suitable for incident reporting within aviation, given that incidents are not generally investigated and interviewing all aircrew involved in incidents would be too costly and logistically infeasible given the large number of incidents that occur. Incident information in the aviation domain is generated using handwritten forms that pilots complete on their own.

Schematic Maps

One possible tool for improving the information recalled in aviation incident reports is a schematic map. Schematic maps use nodes (boxes) and links to spatially represent experts' mental model of a domain or process (Dansereau, 1978; Brooks & Dansereau, 1983). Through visual representation, schematic maps depict the system's main components and interrelationships, as well as the causes and consequences that various inputs have on system performance. Users of these maps fill in the boxes with the details of a particular event as a means for facilitating recall and understanding the factors that influenced their behavior.

Potentially, schematic organizers could be employed as a knowledge elicitation tool for pilots to use prior to writing an incident narrative. Since schematic maps represent expert models of the domain and provide a form that pilots can complete by themselves, schematic maps essentially eliminate the need for a one-on-one interview by an expert. As such, schematic maps have the potential for wide spread application within the aviation industry not afforded by traditional knowledge elicitation methodologies.

Critical Event Reporting Tool

The Critical Event Reporting Tool (CERT, see Figure 1) has been developed to improve the type of information reported within the narrative portion of aviation incident reports. Potentially, CERT provides a structure that prompts or cues the recall of important event information. In addition, the spatial layout of the form may help highlight the interrelationships among factors and the consequences of operator actions in response to the critical event. Empty boxes within the organizer may also provide feedback to the users with respect to gaps in their recollection or description of the incident. Finally, the generic structure the CERT may prove beneficial in its application across incidents or domains.

Nonetheless, such a framework as CERT may also have potential drawbacks for users. For example, this type of schematic organizer may be too inflexible to capture all of the information that users need to report. In addition, the static nature of the CERT form may stagnate the reporting of dynamic, iterative events that often occur in the aviation domain. Finally, the complexity of the form or its "busyness" may also prove too complex or difficult for novice users to employ effectively.

METHOD

The purpose of this study was to empirically evaluate CERT in order to assess its potential strengths and weaknesses as a pre-organizer for event reporting. Specifically, general aviation pilots (n = 34) were exposed to identical mechanical failures during a simulated cross-country flight and were then asked to complete an ASRS-type report of the event. Half of the pilots (CERT group, n = 17) received a brief description and blank copy of CERT which they were allowed to study for 5 min. They were then provided a completed example of CERT mapped to a fictitious automobile accident (5 min). Next they read an example essay/incident report which corresponded to the same automobile accident (5 min). The other half of the pilots (Control group, n = 17) completed a short distracter task (10 min) that involved reading about the ASRS system. They were then provided the essay report of the fictitious automobile accident and allowed to study it for 5 min. Pilots in both groups then began the task of incident reporting to describe the events they had experienced during

the simulated flight. Hence, both groups began the process of recall after a 15 min delay. Pilots in the CERT group were given 15 minutes to complete the form prior to writing the essay, whereas pilots in the control group wrote the essay

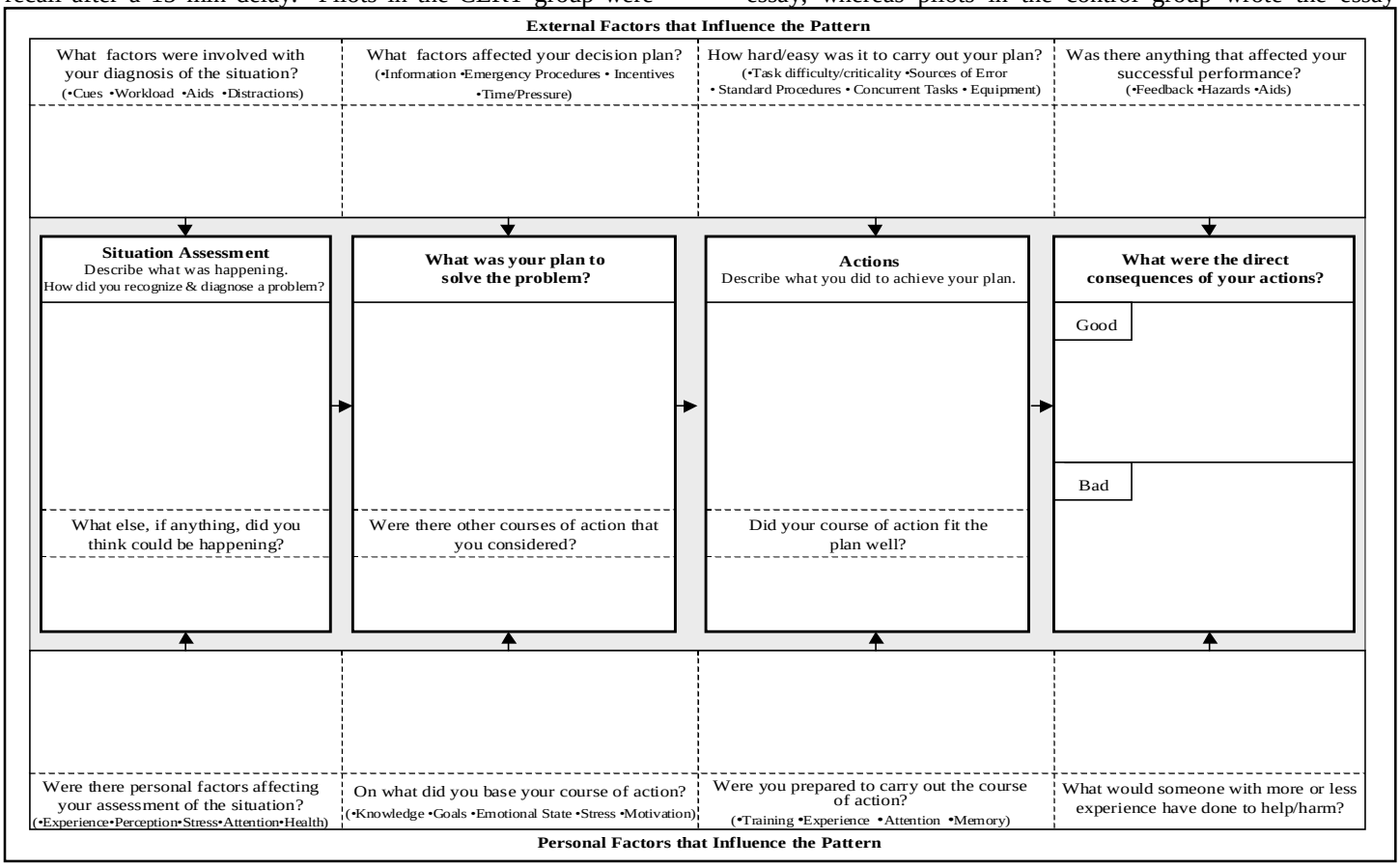


Figure 1. The Critical Event Recall Tool (CERT).

only. Pilots in both groups were given as much time as needed to complete the essays. Following completion of the essay, pilots in both groups were asked to rate the content and format of their essays using a set of eleven 7-point Likert scale questions. The CERT group was administered an extra questionnaire that asked them to rate CERT in terms of its format and effectiveness as a pre-organizer using a set of eleven 7-point Likert scale questions.

RESULTS

Subjective Evaluations of CERT

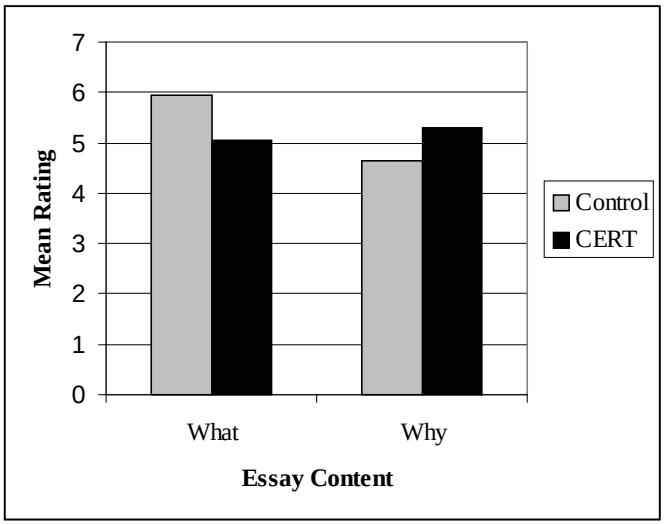
All evaluations of the CERT were above neutral using the 7-point Likert scale format (with higher scores reflecting ratings that are more positive). In general, pilots in the CERT group felt that the CERT form aided them in their recall of the incident (M = 4.65, SD = 1.22) $t(16) = 2.18$, $p<.05$, and was flexible enough to adapt to their specific requirements (M = 4.9, SD = 1.41), $t(16) = 2.58$, $p<.05$. They also found the form helpful in highlighting where important information may have been left out when recalling the event (M = 5.12, SD = .86), $t(16) = 5.37$, $p<.01$, and generally felt it had adequate size and space for writing in the boxes (M = 5.4, SD = 1.94), $t(16) = 3.0$, $p<.01$. In addition, they felt that the form was not too technical to understand (M = 6.0, SD = 1.22), $t(16) = 6.73$, $p<.01$, nor too generic to meet their needs (M = 5.6, SD = 1.3), $t(16) = 5.13$, $p<.01$.

Subjective Evaluations of Essays

Pilots in both groups rated the content and quality of their essays using a set of eleven, 7-point Likert scale questions, with higher scores reflecting more positive ratings. These items were analyzed using a logistic regression procedure in which treatment group (CERT vs. Control) was the dependent variable and responses on the questionnaire items were the predictor variables. The results of this analysis revealed a significant logistic regression function ($X^2 = 11.83$, $p < .003$). In particular, two questionnaire items combined to significantly discriminate between groups. One item pertained to pilots' impressions of how well they felt their essays described what happen during the incident, whereas the second item pertained to the impressions of how well their essays described why the events happened. As seen in Figure 2, pilots in the control group (essay only) felt that they were better at relaying to the reader a sense of what happened during the incident as opposed to those who used the CERT.

In contrast, pilots who used the CERT felt they were better at relaying why they performed the way they did, as opposed to those who wrote the essay only.

Figure 2. Subjective evaluations of essay performance.



Objective Evaluations of Essays

Quantitative Analysis. The amount of time to complete the essays for pilots in the CERT group (M = 15.82 min., SD = 4.17) did not differ significantly from completion times of pilots in the control group (M = 15.41 min., SD = 6.12). There was also no significant difference in the number of words used to write the essay; the control group with a mean of 284 (SD = 98.74) words per essay, and the CERT group with a mean of 276 (SD = 88.14) words per essay.

Qualitative Analysis. The content of pilots' essays was further analyzed by categorizing statements into one of three broad categories, including *what* happened (i.e., descriptive statements), *why* something happened (i.e., analytical statements), and *context* statements (i.e., framing statements). These scores were then analyzed using a logistic regression procedure in which treatment group (CERT vs. Control) was the dependent variable and statement type was the predictor variable. The results of this analysis revealed a significant logistic regression function (X^2 = 4.674, p = .031, df=1). In particular, only one statement type significantly discriminated between groups. Specifically, analytical, or *why*, statements constituted a larger percentage of participant's essays in the CERT group (M = 20%) than in the control group (M = 13%).

DISCUSSION

The Critical Event Recall Tool (CERT) appears to have notable potential as an aid in critical incident reporting. Pilots' evaluations of the value and utility of CERT was overall very positive. CERT was judged to aid the recall of the incident; was deemed flexible; and was viewed as being helpful in highlighting where important information may have been left out in recollection of the event. In addition, the CERT was not seen as being too confusing or restrictive in nature, nor did the pilots who participated in this study view it as being too generic to fit the event experience.

Subjective evaluations of the essays revealed that the control group felt as though they had included more descriptive content in their essays than did the CERT group. In contrast, however, pilots who used CERT reported feeling that the schematic organizer improved the analytic content of their essays. Objective content analysis confirmed that these evaluations by the participants were generally correct. In particular, pilots who used the CERT did, on average, have more analytical statements about why things had occurred during the incident. However, it should also be noted that, contrary to subjective estimates, it did not appear that the control group included more descriptive statements in their accounts than the CERT group.

The results of the present study corroborate previous findings that free-format reporting generally results in essays that focus primarily on what happened rather than why an event occurred. The use of a theoretically based tool, such as CERT, in addition to free format reporting (as with the ASRS system) can be used to improve the analytical content of these reports without sacrificing the essential specifics of *what* happened also. It should be noted that participants in this study received only brief training on how to use CERT and were only allowed to use the tool for a maximum of 15 min prior to writing their essays. This procedure was implemented in order to address issues related to the real-world application of the tool, which would generally be constrained by pilots' willingness to spend time preparing to write the report. Additional research is needed, however, to explore the applicability of CERT to other types of events and situations within the aviation domain.

ACKNOWLEDGEMENT AND DISCLAIMER

The authors thank Bill Geibel and Donald Talleur of the Aviation Research Lab at the University of Illinois for their help with this project. Their assistance and enthusiasm made data collection for this study nearly effortless. This material is based upon work supported by the Federal Aviation Administration under Award No. DTFA 99-G-006. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Federal Aviation Administration.

REFERENCES

Brooks, L. & Dansereau, D. (1983). Effects of structural schema training and text organization on expository prose processing. *Journal of Educational Psychology*, 75.

Connell, L. (1999). Incident reporting: The NASA aviation safety reporting system. *GSE Today*.

Dansereau, D. (1978). The development of learning strategy curriculum. In H.F. O'Neil, Jr. (Ed.), *Learning strategies*. New York: Academic Press.

Diaper, D. (1989). *Knowledge elicitation: Principles, techniques and applications*. NY: John Wiley and Sons.

Ericsson, K. & Simon, H. (1980). Verbal reports as data. *Psychological Review*, 87(3).

Flanagan, J. (1954). The critical incident technique. *Psychological Bulletin*, 54(4).

Geiselman, R. & Fisher, R. (1989). The cognitive interview technique for victims and witnesses of crime. In D.C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence*. New York: Springer.

Hall, E., Gott, S., & Pokomy, R. (1995). *A procedural guide to cognitive task analysis: The PARI methodology* (Technical Report AL/HR-TR-1995-0108). Brooks AFB, TX: Armstrong Laboratory.

Heinrich, H. W. (1959). *Industrial accident prevention: Ascientific approach* (4th ed.). New York: McGraw-Hill.

Hiland, D. & Dzieszkowski, P. (1984). Hypnosis in the investigation of aviation accidents. *Aviation, Space and Environmental Medicine*, December.

Kirwan, B. & Ainsworth, L. (1992). *A guide to task analysis.* Washington, DC: Taylor & Francis.

Klein, G., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man & Cybernetics*, 19(3).

Klein, G. (1993). *Naturalistic decision making: Implications for design.* Ohio: CSERIAC SOAR 93-1.

Loftus, E., Greene E. & Doyle, J. (1989). The psychology of eyewitness testimony. In D.C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence*. New York: Springer.

Memon, A. & Bull, R. (1991). The cognitive interview: Its origins, empirical support, evaluation and practical implications. *Journal of Community & Applied Social Psychology*, 1.

Militello, L., & Hutton, R. (1998). Applied cognitive task analysis: A practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, 41.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5.

Wiegmann, D. A., Danserau, D. F., Skaggs, L. P. & Gordon, N. B. (1992). Comparing techniques for processing schematic maps. *Southwestern Psychological Association*.

Wiegmann, N. G. (1992). Schematic organizers and cooperative review: Impact on recall and metacognition.*Master's Thesis: Texas Chrisitian University*.